Affordability Ratio Methodology Webinar

R. 18-07-006

February 27, 2020

Agenda

• Introduction • Goals & Expectations	Wylen Lai
• Methodology Overview	Ankit Jain
• Geographic Information Systems (GIS) • Q&A on GIS	Jeremy Ho
• Regression Model • Q&A on Regression Model	Ankit Jain
• AR Results • Q&A on AR Results	Jefferson Hancock
• Q&A, Next Steps	Staff

Webinar Goals

- Provide high level overview of AR calculation methodology in the Revised Staff Proposal
- Discuss key aspects of the methodology
  - Geographical Information System (GIS) intersect of service territory and census geography
  - Regression model
  - AR results
- Q&A expectations

Webinar Context

- Revised staff proposal recently submitted with key changes to prior methodology for AR calculation
  - Industry-specific AR in addition to bundled AR
  - Electricity and gas as two separate utilities instead of combined “energy”
  - Housing costs determined by regression analysis (function of household income and size)
  - Developed aggregation technique so that average AR can be reported for any desired geographic scale

AR Calculation Methodology

Essential utility service charges (water, electric, gas, and communications)

Household income minus housing cost

Combined Bundle: $AR_{i,total} = \frac{W+E+G+C}{i-H}$

Electric: $AR_{i,E} = \frac{E}{i-(H+W+G+C)}$

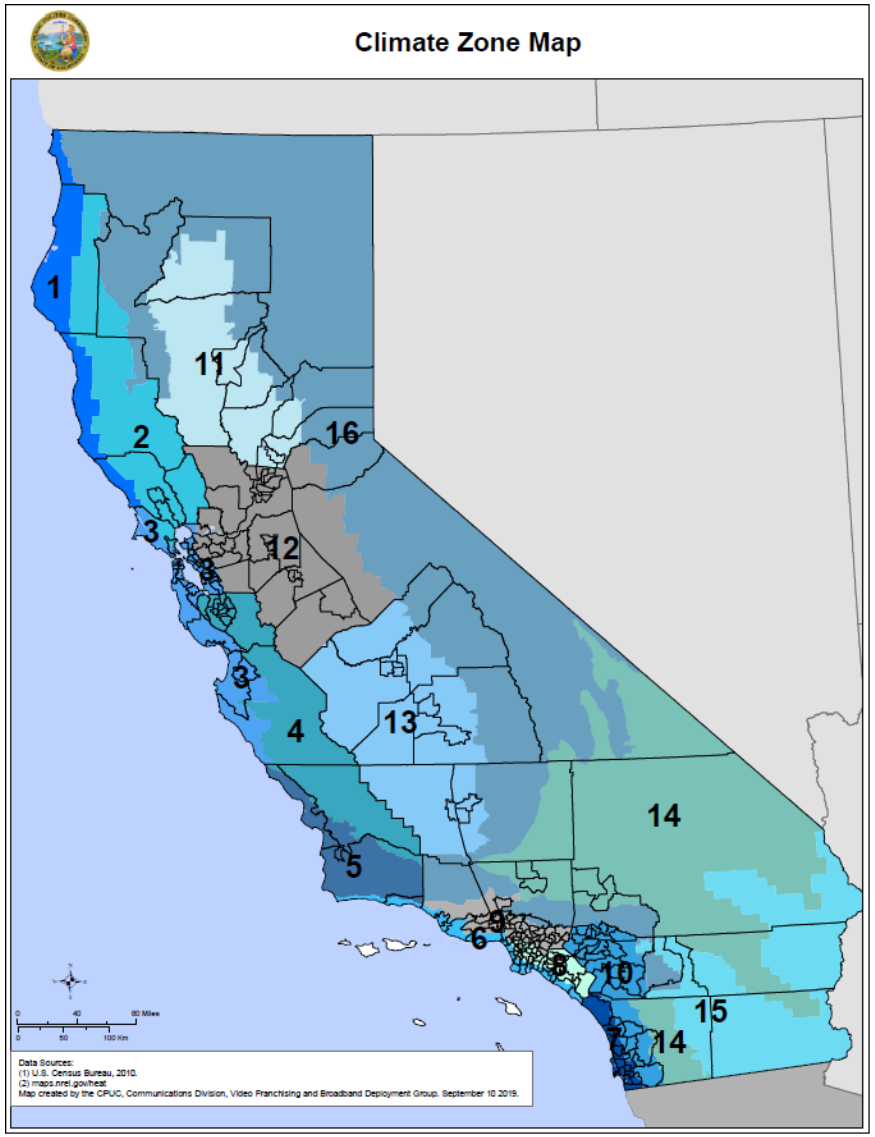
Gas: $AR_{i,G} = \frac{G}{i-(H+W+E+C)}$

Communications: $AR_{i,C} = \frac{C}{i-(H+W+E+G)}$

Water: $AR_{i,W} = \frac{W}{i-(H+E+G+C)}$

- Calculation of cost at essential usage level
- Calculation of household income and housing costs
- Calculation of AR at service territory/block intersect scale
  - Top coding
- Aggregation

Boundary Mismatch



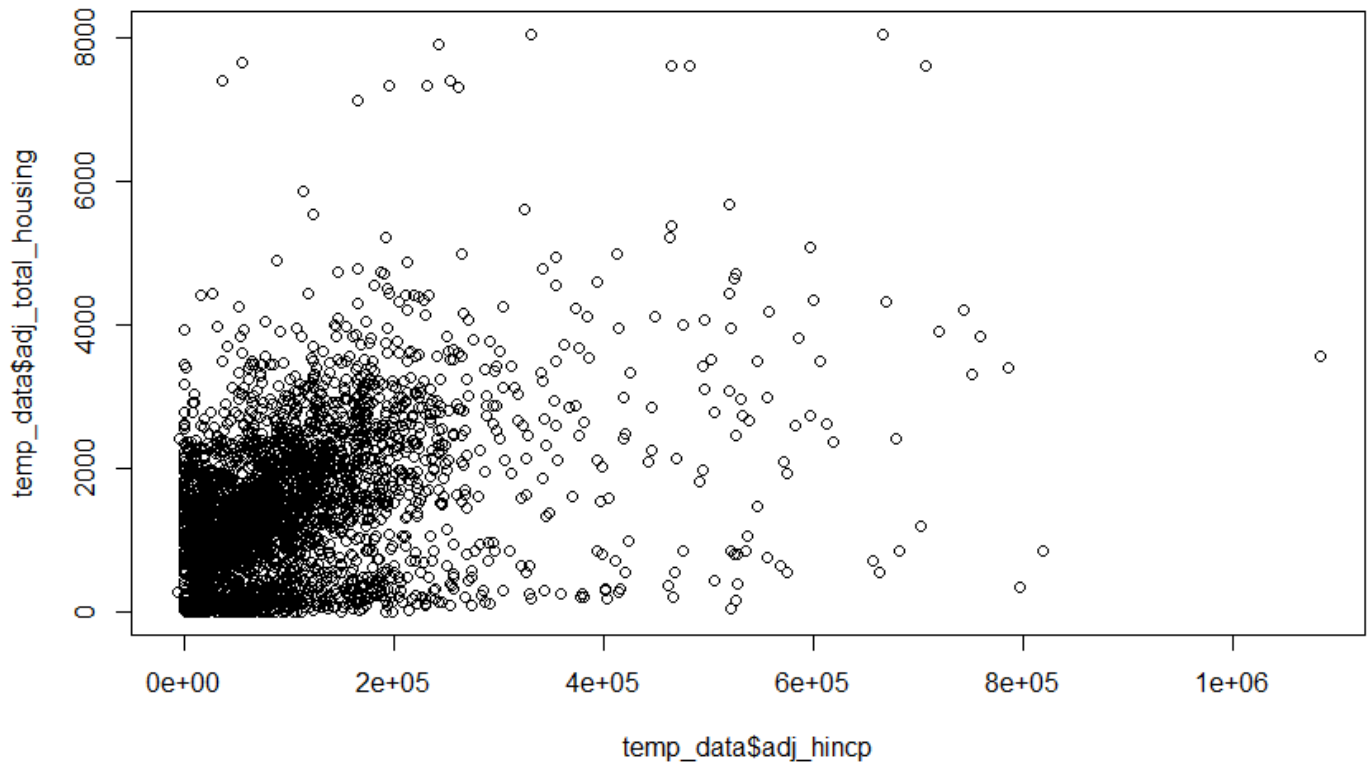
AR Calculation Methodology

- Calculation of cost at essential usage level
- Calculation of household income and housing costs
- Calculation of AR at service territory/block intersect scale
- Aggregation

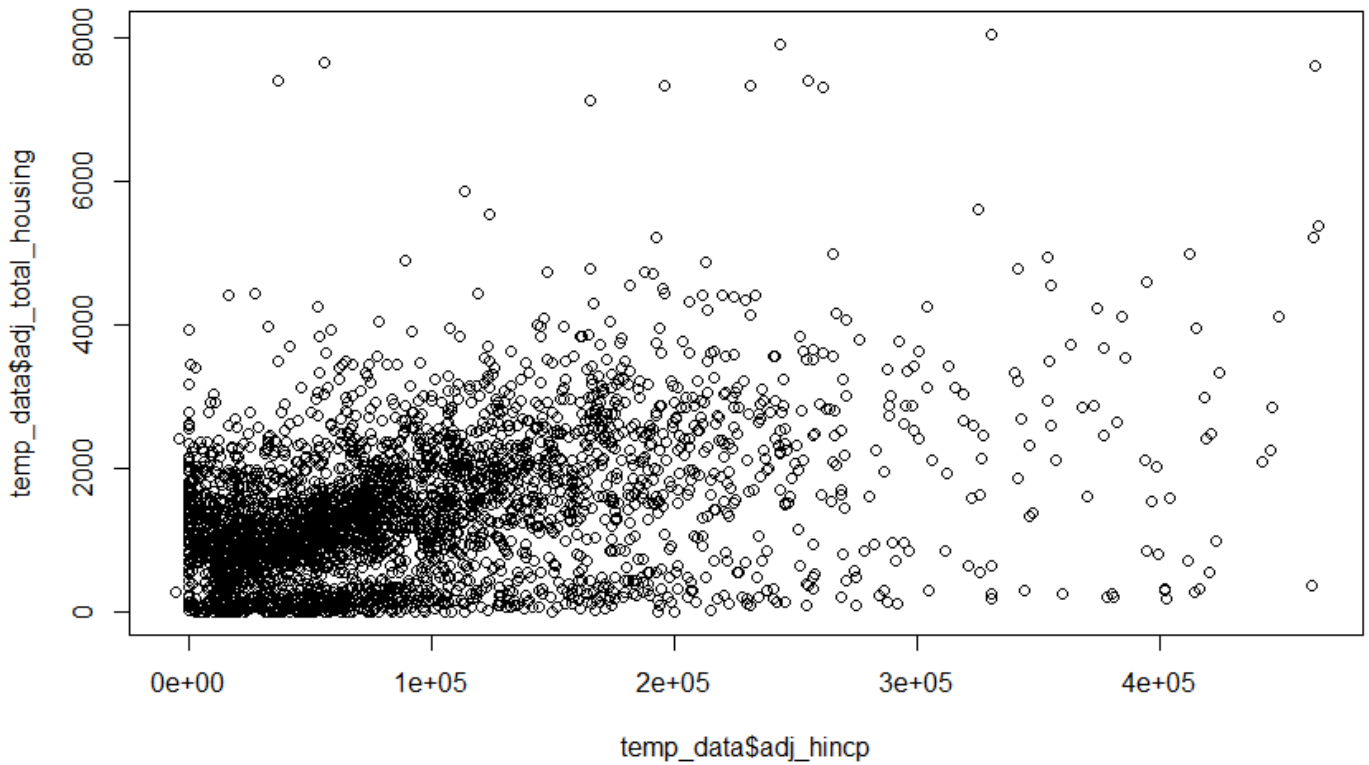
Regression Analysis Context

- Previous methodology for estimating housing cost for a given income level:
  - Would have required assignment of utility provider to each household
  - Arbitrary size of income band
  - Result would be sensitive to the skewedness of income distribution
  - Would not address the wide range of housing costs since the final reported number would still be an average of the sampled households
- Goals of updated approach:
  - Allow for calculation of AR at service territory/block intersect scale (to avoid utility assignment and allow for aggregation to any desired geographic scale)
  - Provide point estimate of income and corresponding housing cost for given income percentile based on observed data
    - Acknowledge that housing costs exhibit a high degree of variance, and the results for an individual household may deviate from point estimate

Example plot of housing cost vs. income (PUMA 11300)



Same plot with high income households removed (> 5x weighted mean income)



Regression models tested

- Linear models
  - Housing Cost = a + b*Income + error
  - Housing Cost = a + b*Income + c*Household Size + error
  - Housing Cost = a + b*Income + c*Household Size + d*(Income*Household Size) + error
- Square root functional form models (to account for overestimation at high income levels)
  - Housing Cost = a + b*$\sqrt{Income}$ + error
  - Housing Cost = a + b*$\sqrt{Income}$ + c*Household Size + error
- Repeated tests on same models without high income households (> 5x weighted mean income for a given PUMA)
- All regressions run on each individual PUMA

Regression results

- Models predict modest coefficients for income and household size
  - Predictions vary a modest amount as a function of income (~$50/month increase for every $10k of income) and household size (~$100/month increase for each additional person in household), but are slightly different for each model
- All models produce small p-values for all coefficients
  - Income and household size are strong predictors of housing cost, and work well in all the models presented here
- All models have small R-squared values (~0.15 to 0.2)
  - Although income and household size have strong positive correlation with housing cost, there is a considerable amount of variation in housing cost that is not explained by these two predictors alone

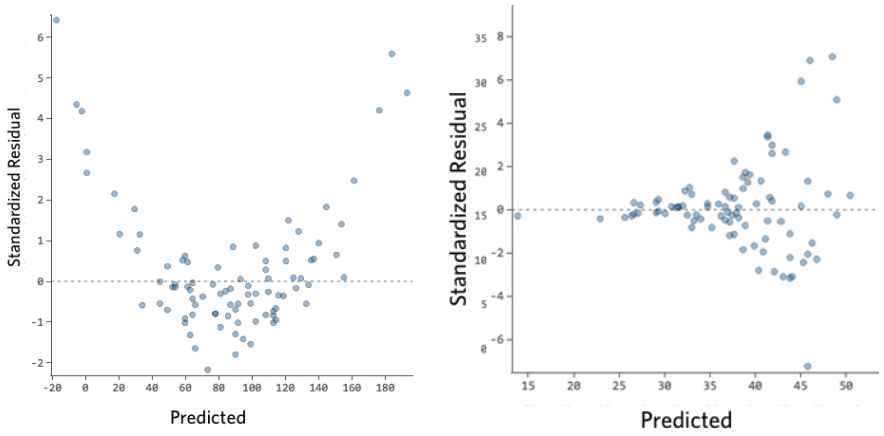
R^2 Outputs

Model	Average R^2 Across All PUMAs	
	All Customers	No High Income (> 5x mean income)
Housing Cost = a + b*Income	0.146	0.158
Housing Cost = a + b*Income + c*Household Size	0.173	0.181
Housing Cost = a + b*Income + c*Household Size + d*(Income*Household Size)	0.176	0.184
Housing Cost = a + b*√Income	0.167	0.164
Housing Cost = a + b*√Income + c*Household Size	0.187	0.183

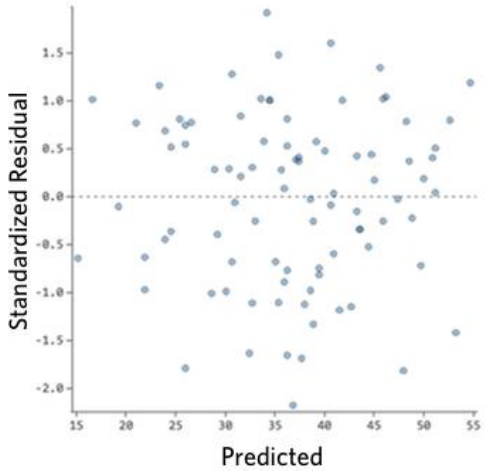
Interpreting residual plots

- Ideally, residuals (observed – predicted value) are:
  - Randomly distributed around zero (predictions are not biased one way or the other)
  - Do not show a pattern as a function of income (model does not do a worse job predicting at one income level vs another)

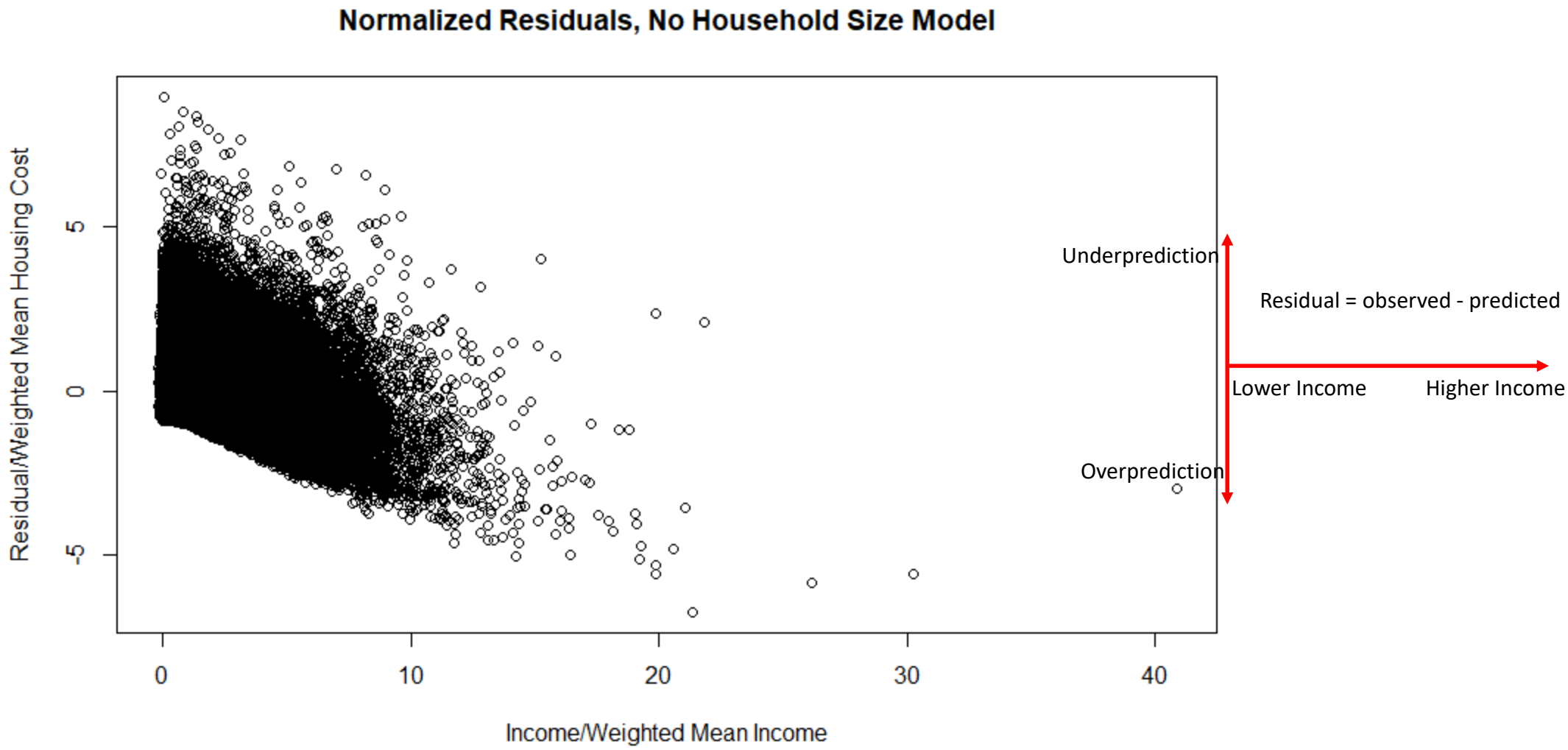
Residual Plots for Bad Models



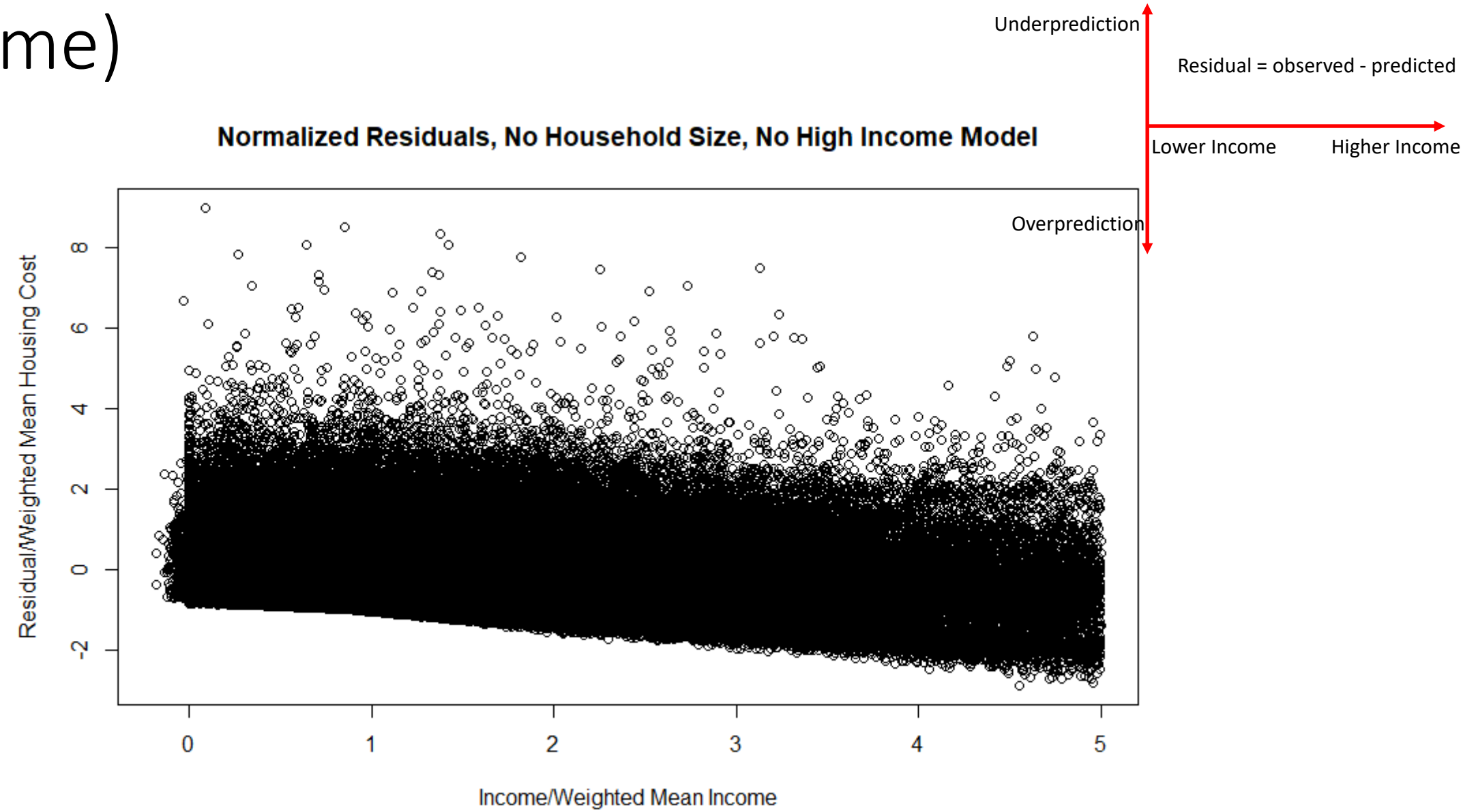
Residual Plot for Better Model



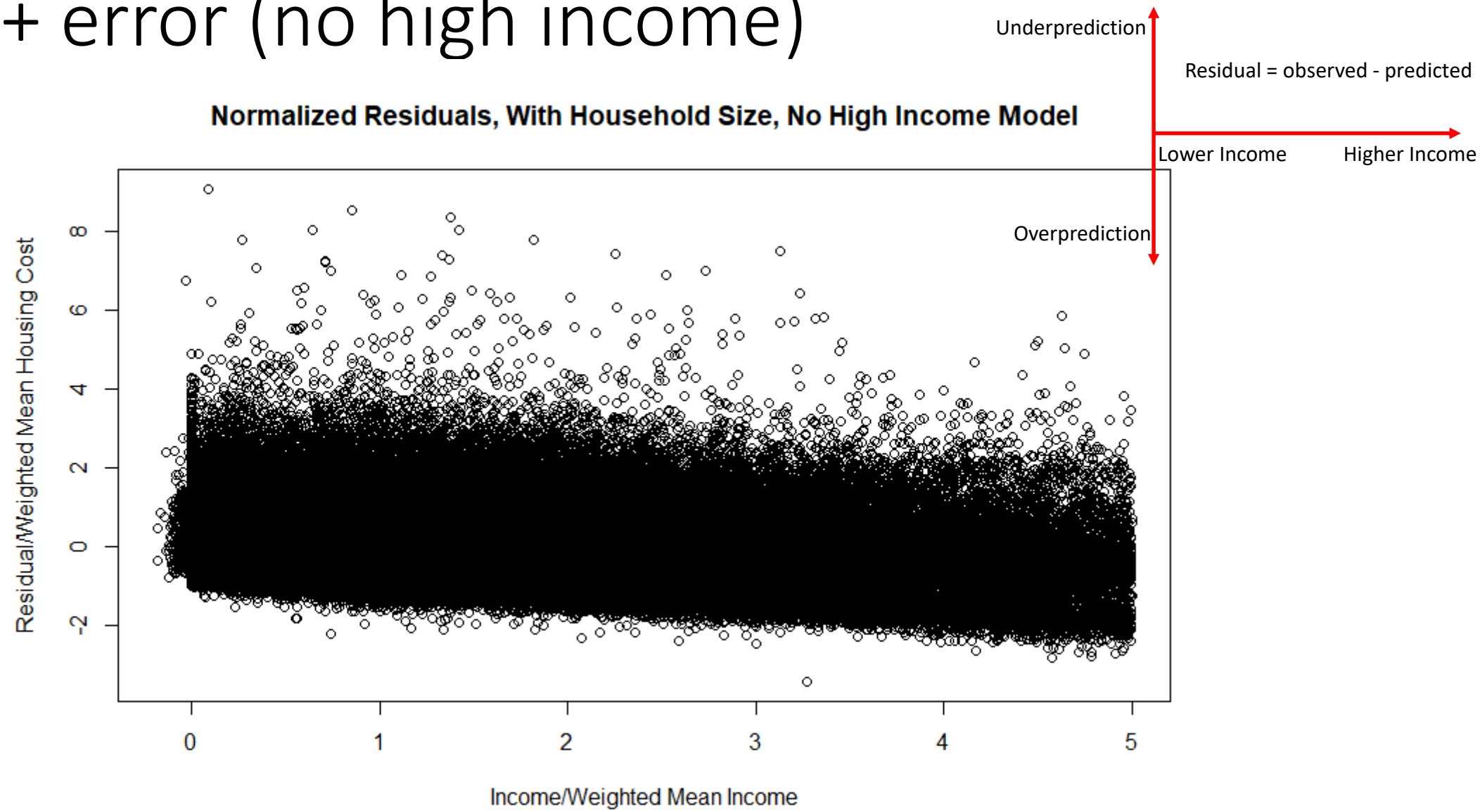
Housing Cost = a + b*Income + error



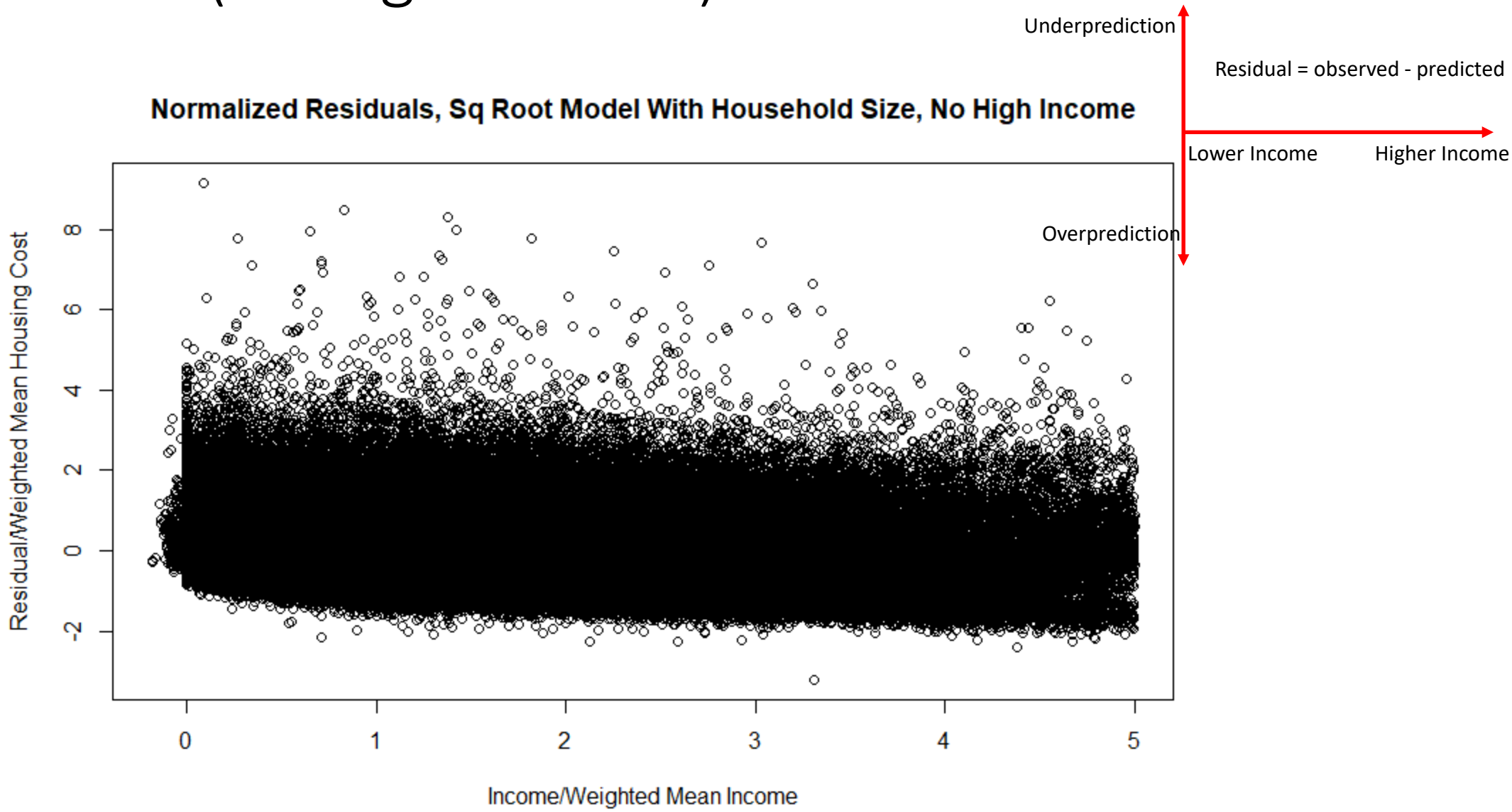
Housing Cost = a + b*Income + error (no high income)



Housing Cost = a + b*Income + c*Household Size + error (no high income)



Housing Cost = a + b*$\sqrt{Income}$ + c*Household Size + error (no high income)

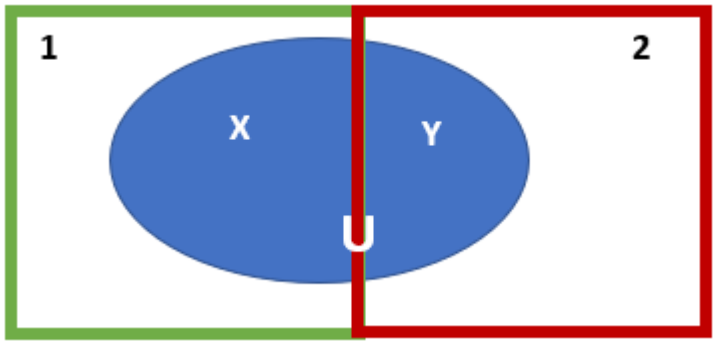


AR Calculation Methodology

- Calculation of cost at essential usage level
- Calculation of household income and housing costs
- Calculation of AR at service territory/block intersect scale
- Aggregation

Aggregation Process

- Due to misalignment of utility service boundaries and PUMA boundaries, AR is calculated at scale at which there are unique values of AR determinants and then aggregated into a weighted average:
  - Intersect utility service and PUMA boundaries to identify areas with unique combination of utility charges and income/housing costs
  - Calculating AR value for each of these areas
  - Aggregate individual ARs across service territory using a weighting mechanism that accounts for area and population



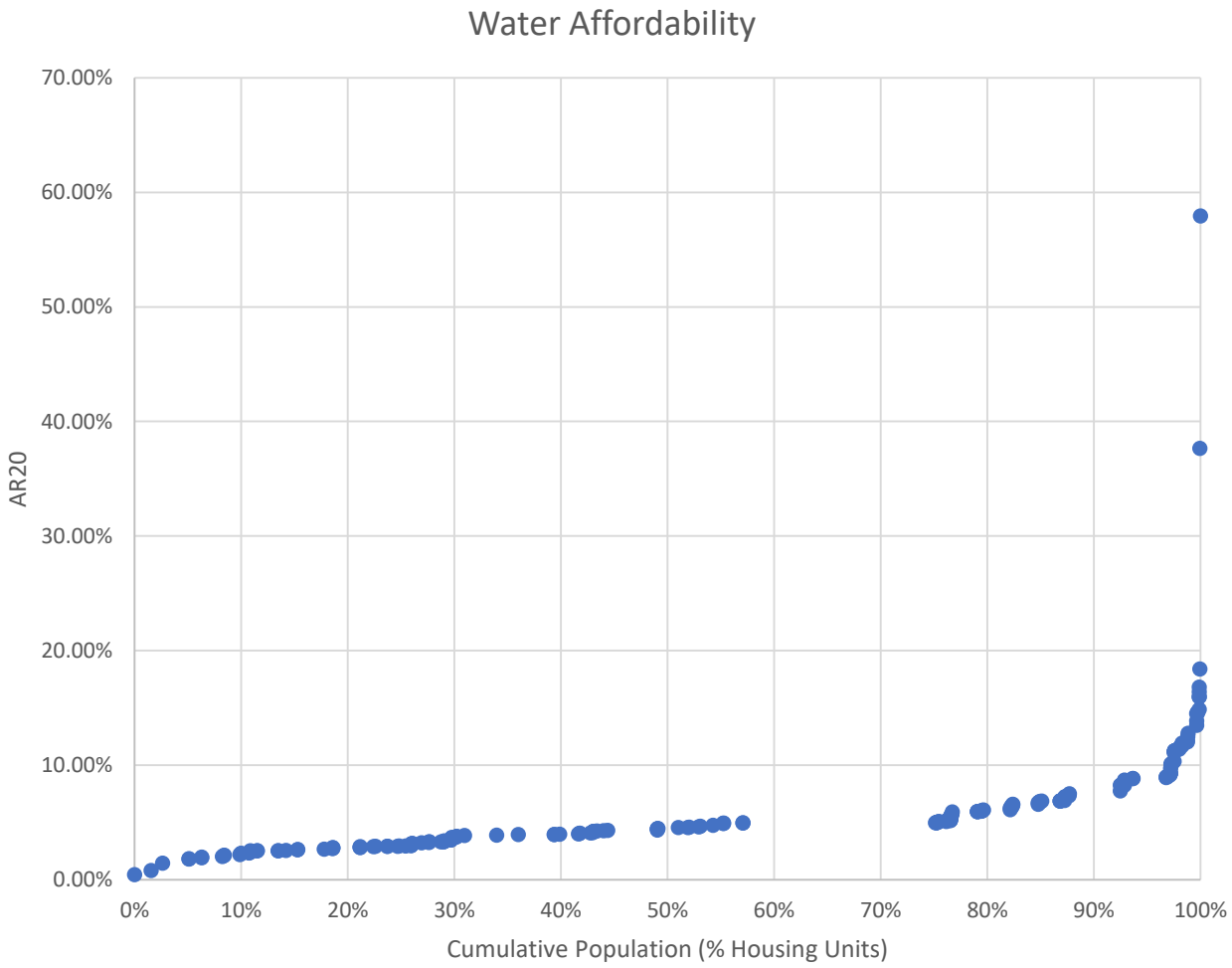
$$W_x = \frac{A_x P_1}{A_x P_1 + A_y P_2}$$

$$W_y = \frac{A_y P_2}{A_x P_1 + A_y P_2}$$

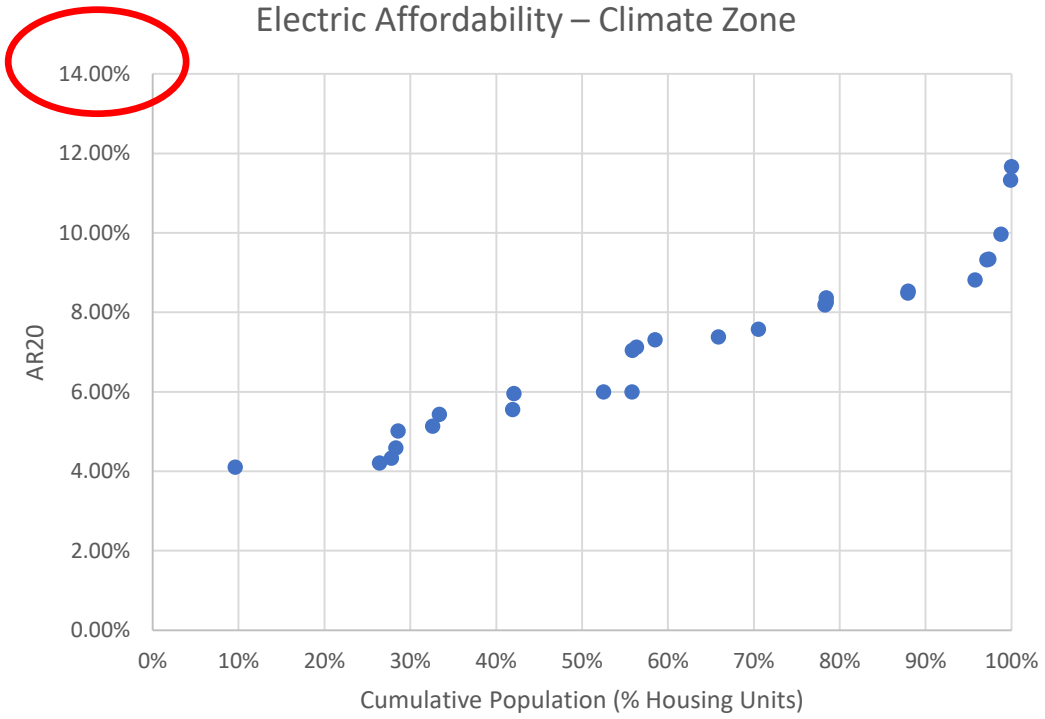
U = Utility Service Area
X = Portion of U in PUMA 1
Y = Portion of U in PUMA 2

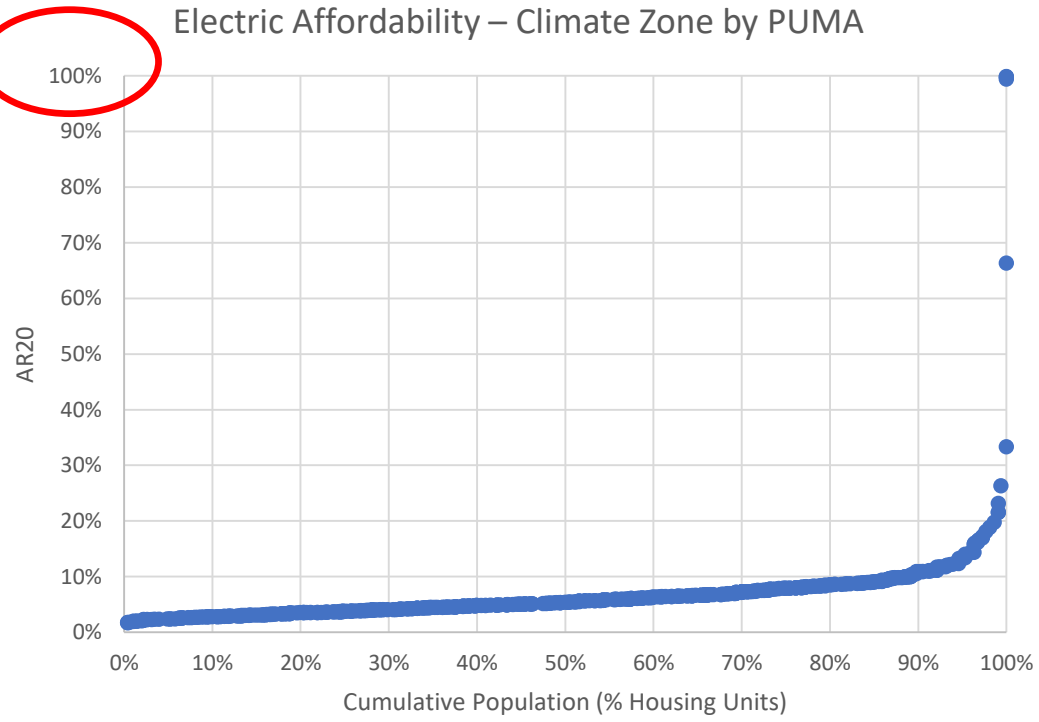
Affordability Distribution: Water

- Each point is a different water system
  - Larger x-axis gaps represent more people served within the same system
- 75% of population has an AR20 value below 5%
- 97% of the population has an AR20 value below 10%
- Affordability concerns especially pronounced for a relatively small segment of the population, for whom affordability is a major issue



Electric Affordability



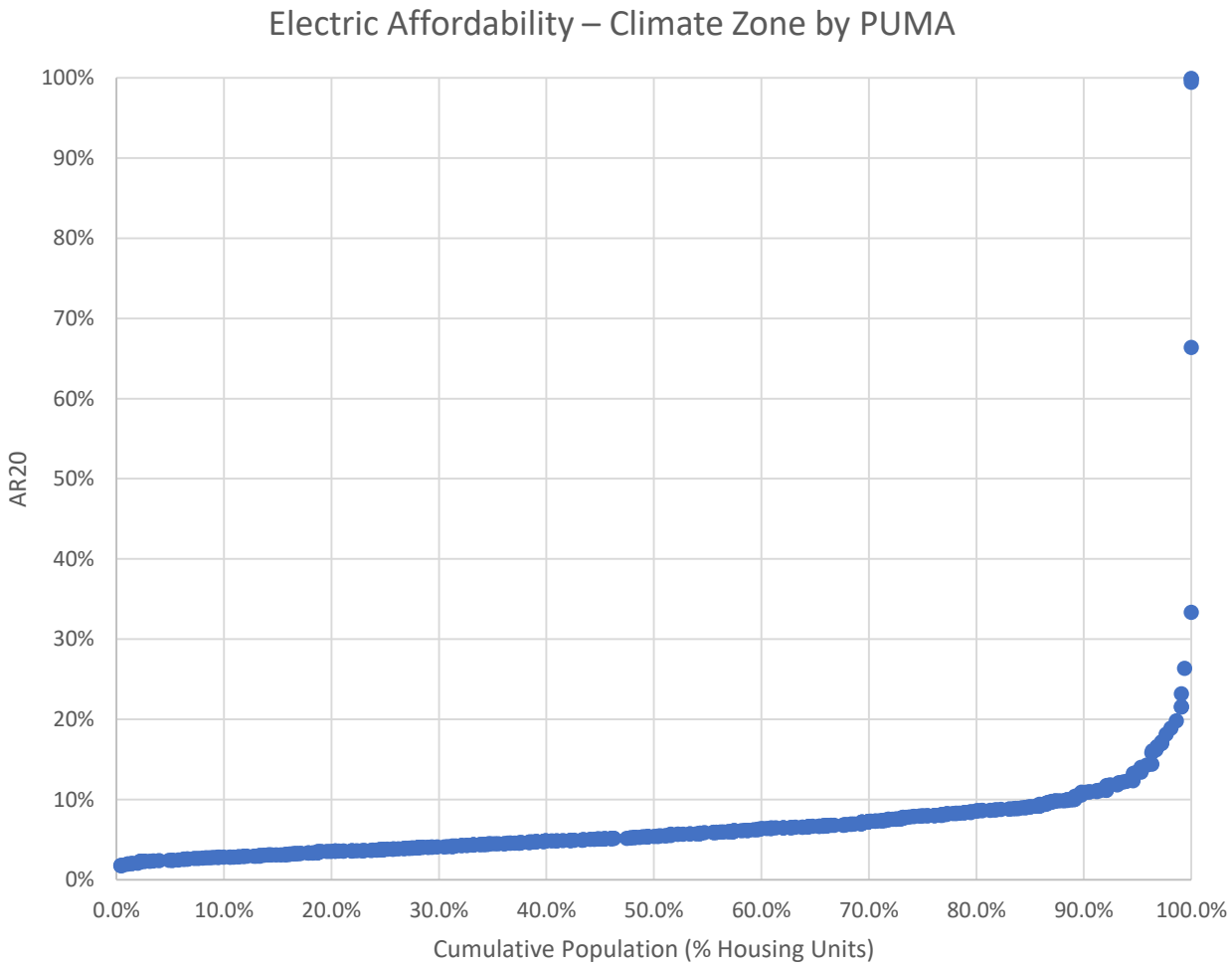


- Climate zone-level analysis obscures regional variation
- Understates unaffordability at highest levels
- Can still be useful for comparisons between climate zones, “big picture” analysis

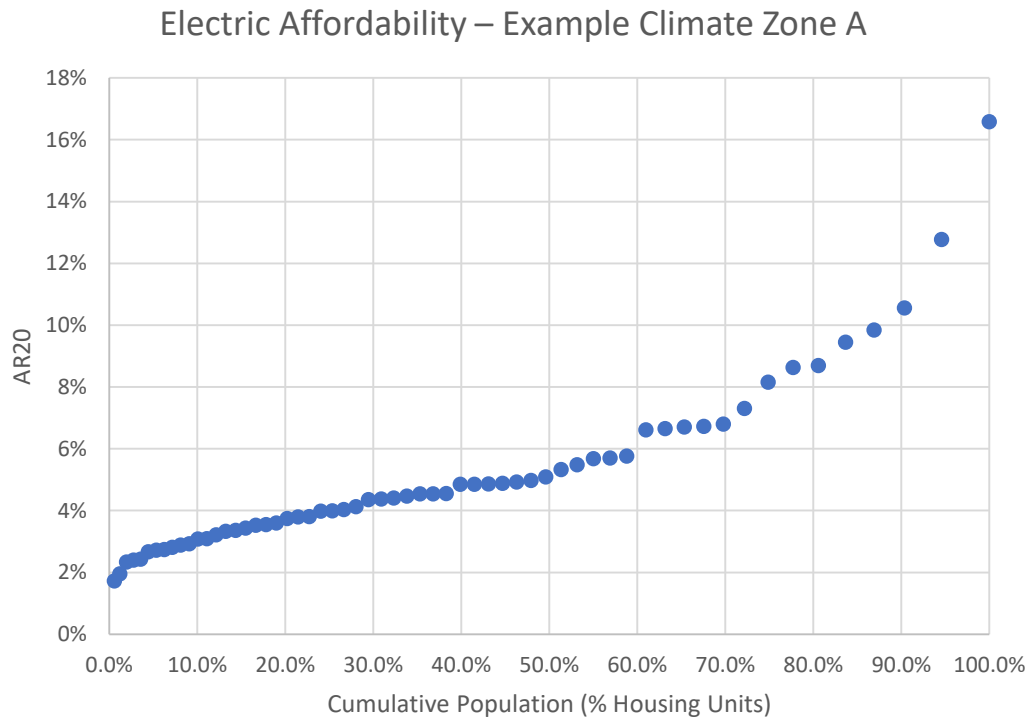
- Each point is a combination of electric climate zone and PUMA
  - So, each PUMA in a given climate zone is a separate data point
  - If multiple climate zones exist within the same PUMA, they are also shown separately
  - 500 combinations total

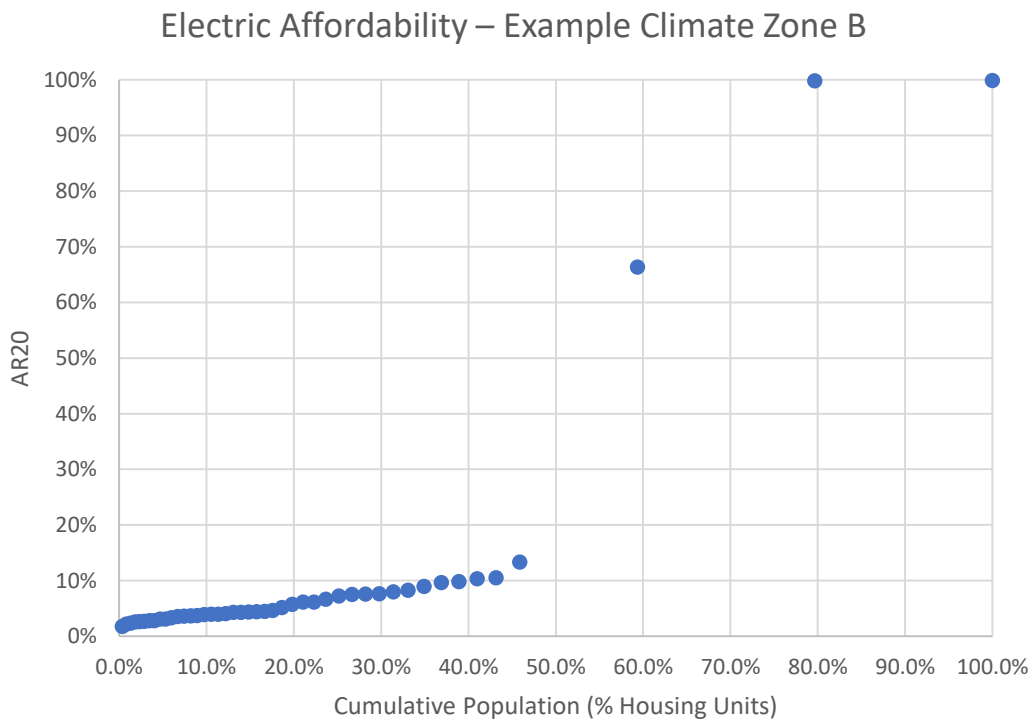
Affordability Distribution: Electric by PUMA

- We can aggregate to the scale of a utility or climate zone, but often we want to see variation within a climate zone
- Question: how affordable is the same rate structure across different areas?
- Around 90% of housing units have AR20s of 10% or below
- As with water, very high unaffordability for a very small proportion of the population



Electric Affordability – Comparison Within & Between Climate Zones

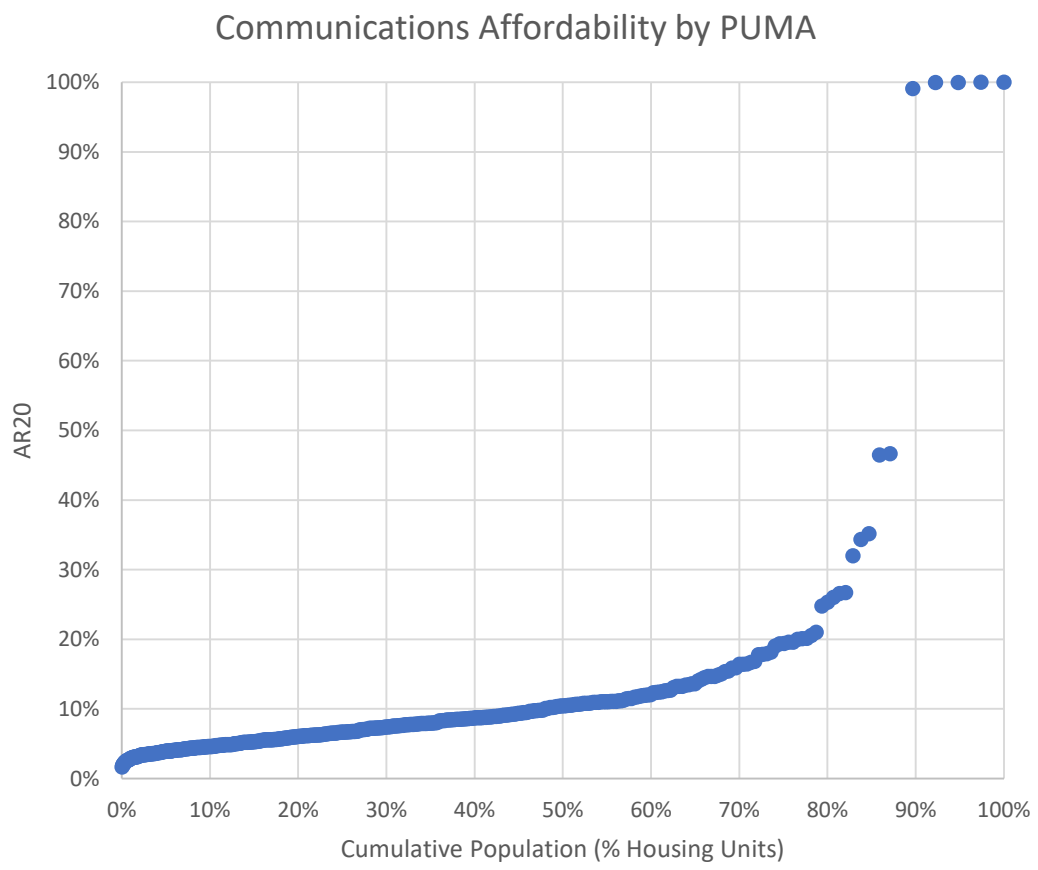


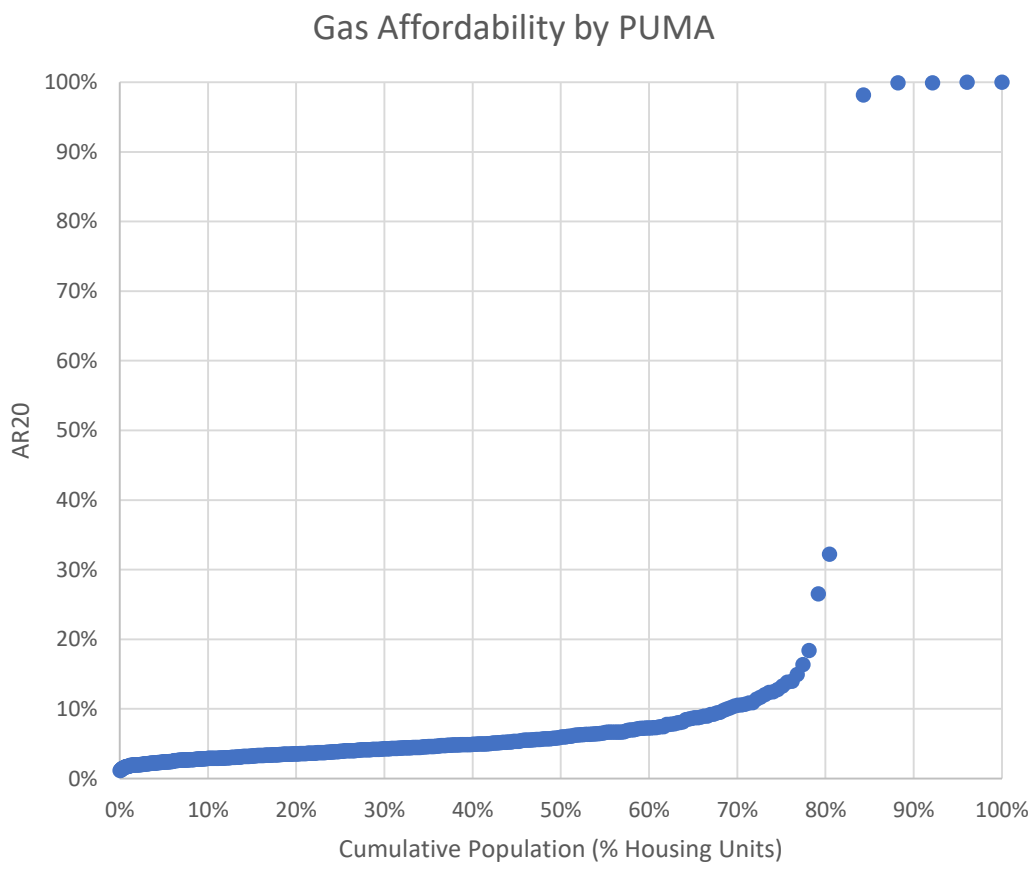


- 90% of HUs have AR20 <10%
- Relatively low maximum of ~16%

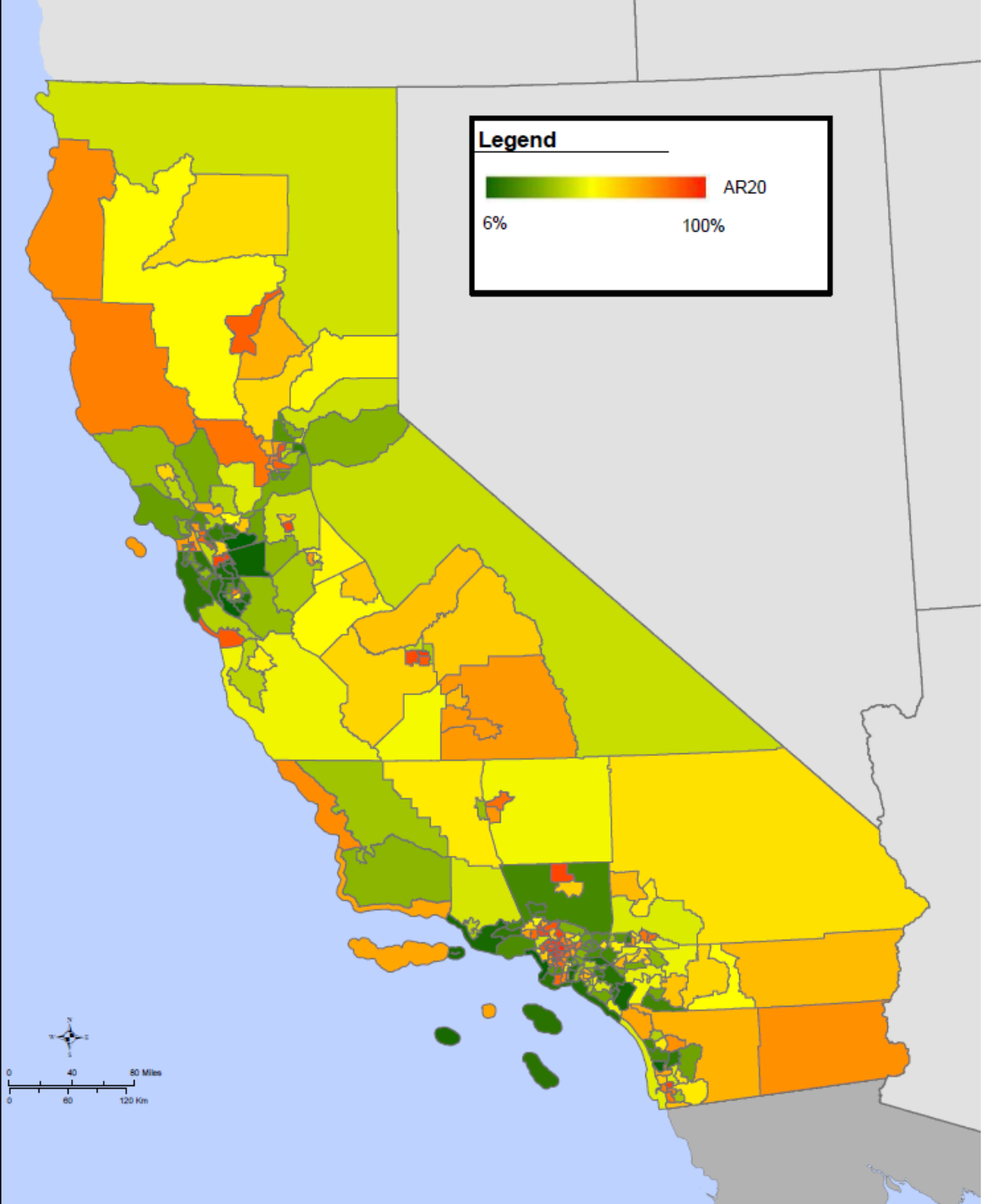
- ~50% of population has AR of <13%, but remaining 50% has very high ARs
- Maximum AR of 100% (topcoded)

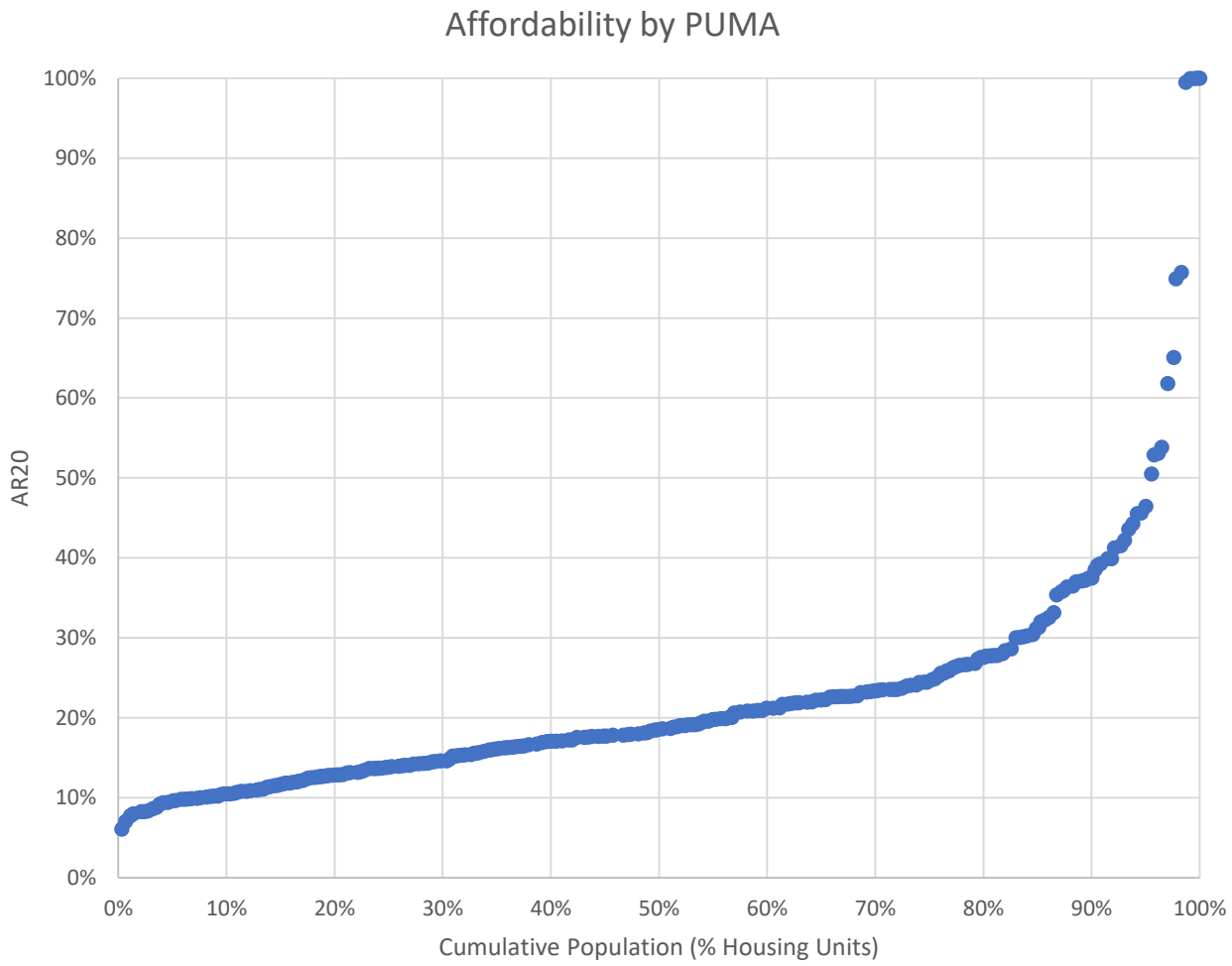
More Affordability Distributions





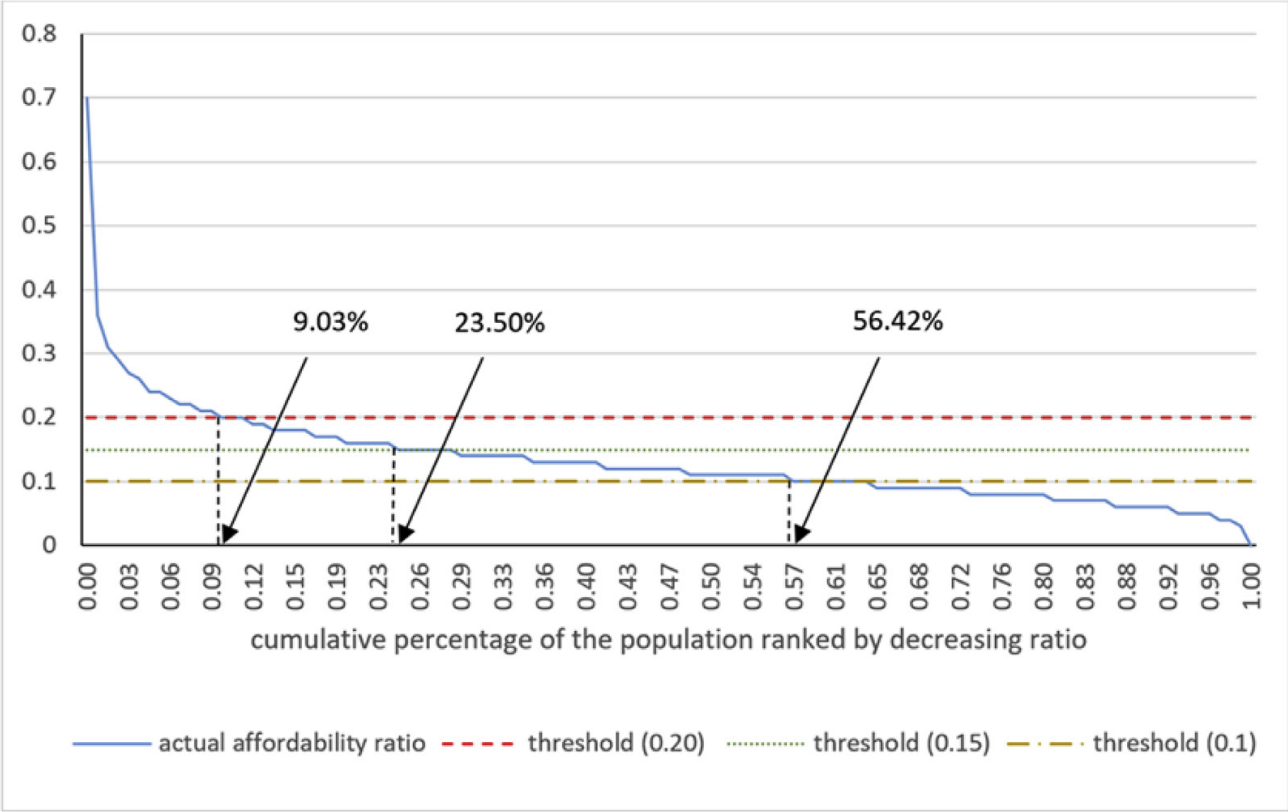
Combined Affordability – PUMA Scale





Next Steps

- Annual report: how else should this data be presented?
- Compare to indicators of hardship
  - Disconnections, arrearages, etc.
- Application in proceedings
  - Forecasting
  - Other sources of data
  - Other questions



Source: Martins et al. 2019. "Making ends meet: Actual versus potential joint affordability of utility services," *Utilities Policy,* vol. 56, pp. 120-126. https://doi.org/10.1016/j.jup.2018.12.002

Thank you

www.cpuc.ca.gov/affordability

Jefferson Hancock, Water Division

Jeremy Ho, Water Division

Ankit Jain, Energy Division

Wylen Lai, Communications Division

For inquiries, please direct them to Jefferson Hancock at JHO@cpuc.ca.gov